

GENETIC CHARACTERIZATION OF LANDRACES AND VARIETIES USING FUNCTIONAL MICROSATELLITE MARKERS IN *ELEUSINE CORACANA* (L.)

NAGABHUSHANA KOTRAPPA¹, SHAILAJA HITTALMAN² & KRISHNA VENKATARANGAIAH³

^{1,3}Department of Biotechnology, Jnanasahyadri, Shankaraghatta,

Kuvempu University, Shimoga, India

²Department of Genetics and Plant Breeding, University of Agricultural Sciences,

GKVK, Bengaluru, Karnataka, India

ABSTRACT

About 3515 genomic and EST sequences from the public database were mined in-silico for microsatellite repeats containing di-, tri-, tetra- and Penta-nucleotide repeat motifs and 3334 primer pairs were designed. One hundred primers were synthesized and validated in *Eleusine coracana* species, and the polymorphic microsatellites were used for characterizing 32 local land races and commercial varieties collected from five different districts in Karnataka. The studies were focused on estimating genetic parameters of the markers, genetic diversity and population structure. Functional annotation and gene ontology was assigned using Uniprot database under the categories of molecular function, biological process and cellular component. The mean values of the number of alleles per locus, diversity index and polymorphic information, content were 4.6, 0.5 and 0.43 respectively. The population structure analysis performed under an admixture model with correlated allele frequencies using STRUCTURE, and the principal coordinate analysis indicated four subgroups within the population. While AMOVA indicated significant variation within and among populations, Hardy-Weinberg test indicated no significant deviations in majority of the markers.

KEYWORDS: Finger Millet, Microsatellites, Landrace, Genetic Diversity & Population Structure

Received: Apr 04, 2017; **Accepted:** May 04, 2017; **Published:** May 26, 2017; **Paper Id.:** IJASRJUN201736

INTRODUCTION

Finger millet (*Eleusine coracana* L.) commonly known as Ragi, is one of the extensively grown minor millets in the regions of India and Africa. It ranks sixth in production after wheat, rice, maize, sorghum and Bajra in India. It is considered as one of the most nutritious cereal due to its abundance of high calcium, iron, protein, fiber and other minerals. It is a rich source of non-available carbohydrates with low glycemic index, which is a boon for cardiovascular and diabetic patients. The antioxidant properties resulting from phytates, polyphenols and tannins makes it anti-cancer and anti-aging food crop (Chandra et al., 2016). Besides its nutritional and health benefitting properties, the superior adaptability traits of this crop make it as a storehouse of important genomic resources for crop improvement (Sood et al, 2016). Finger millet is an allopolyploid derived from 'AA' and 'BB' genomes with chromosome number $2n = 4x = 36$. Though the crop originates from Africa, the Indian sub population is reported to be low in genetic/sequence diversity (Salimath et al., 1995). The knowledge of genetic diversity estimated by molecular markers is essential to select genetically distinct germplasm for hybridization programs. In the absence of finger millet genome sequence, attempts have been made to develop microsatellite markers and single nucleotide polymorphism (SNPs) (Gimode et al., 2016). Markers have been employed to study disease resistance, calcium

dynamics, and in association mapping of various agronomic and nutritional traits (Yadav, et al., 2014b; Babu et al., 2014a, c& d). However, like many orphan crops, finger millet lacks genomic resources sufficient for its genetic improvement. Microsatellites (also called as simple sequence repeats: SSRs) are still the cost-effective and robust markers that offer high reproducibility and polymorphism. Currently, there are < 300 informative microsatellites in finger millet and a couple of SNP studies (Sood et al., 2016). While the development of these markers in the native species is expensive, designing EST-derived markers using the public sequences become cost effective and will add to the limited genomic resources available in this species. Besides wild cultivars that serve as gene donors, the local cultivars and landraces with different attributes offer enormous genetic variability and could be used in varietal improvement program. In this study, the publicly available genomic and expressed sequence tag (EST) sequences reported for various traits, including NBS-LRR regions, disease resistance and abiotic stress tolerance from nine Eleusine species were collected and screened for tandem repeats, in order to develop functionally validated EST-derived microsatellites and were employed in studying local cultivars and landraces of finger millet.

MATERIAL AND METHODS

Genomic DNA was isolated from 32 genotypes collected from five regions within Karnataka, ranging from the southern dry region to western Malnad region (Table 1) using a CTAB method (Doyle and Doyle et al., 1990). A set of 3515 publicly available genomic and EST sequences was obtained from NCBI database and the sequences with >80 % 'N' and <50 bp were trimmed off to get high quality ESTs for mining repeat motifs. Repeat prediction was performed using MISA (Thiel et al., 2003) with a moderately stringent parameters of 5 repeats for di-nucleotide repeats and 3 repeats for tri-nucleotide repeats and 2 repeats for tetra-nucleotide repeats besides compound repeats. Primers were designed in SSR containing sequences using the Primer3 tool (Untergasser et al., 2012) with the optimum standards of 50% GC content, melting temperature of 58°C, primer length of 20 bases and the PCR product size ranging from 150 to 300 bases. The annotation of the primers was carried out and grouped under the functional categories of molecular function, biological process and cellular component using UniProt database.

Table 1: List of Finger Millet Genotypes Collected from 5 Different Districts in Karnataka

Location	Local Landraces	Commercial Cultivars
Chintamani	Nalla, Mooruthingala, Gidda, Thella, Indaf	Indaf-5
Pavagada	Kappu, Kunte, Bili, Dappa	-
Shimoga	Karimundaga, Makari, Mooruthingala, Kapputhene, Bili, Kari kaddi, Malali, Dodda, Sannakaddi, Jainumandaga, Thaidala, Kolimade, Bonda and Jagaluru	-
Mandya	-	MR-1, MR-6, KMR-204 and KMR-301
Bengaluru	-	ML-322, ML-31, ML-365 and Indaf-7

Polymerase chain reaction (PCR) was performed by the components of template DNA (20 ng), primers (5pM), 10X PCR buffer, dNTPs (1 µM), MgCl₂ (1.5mM) and Amplitaq Gold *Taq* DNA polymerase (0.3 U) in a reaction volume of 20 µL. The PCR profile with the step-down annealing conditions (T_A) was performed with the initial denaturing of template DNA at 94°C for 5 min, followed by 35 cycles of 94°C for 60 Sec, T_A for 30 Sec and 72°C for 60 Sec, followed by a final extension at 72°C for 8 min. The amplicons were resolved on 2 % high resolution agarose along with the size standards and the allelic data was extracted based on the size. The marker data were analyzed for genetic diversity parameters such as expected heterozygosity (H_e) (~ Diversity Index: DI), Polymorphism Information content (PIC) and Hardy-Weinberg equilibrium test using CERVUS 3.0 (Kalinowskiet al. 2007). Genalex (Peakall and Smouse, 2012) and

STRUCTURE V2 (Pritchard et al., 2000a) were employed for estimating, analysis of molecular variance (AMOVA), principal co-ordinate analysis (PCoA) and population structure respectively.

RESULTS AND DISCUSSIONS

The microsatellite analysis of 3515 genomic and EST sequences indicated the presence of tandem repeats in 2602 sequences, resulting in the detection of 4530 SSRs. The higher number of SSRs was possible due to the presence of multiple repeat motifs in 1789 sequences. A total of 3334 primers were designed for DI, Tri, tetra and pentarepeat motifs along with the compound formats. The repeat prediction using MISA indicated 16%, 42%, 3% and 39% of di- tri-, tetra- and Penta-base repeats respectively. Unlike the current reports on prevalence of more di-base repeats, our study indicates more of tri-base repeats, which could be due to the inclusion of repeat-enriched nucleotide sequences in SSR discovery. Normally, the presence of Penta-repeats is less in finger millet (Babu et al., 2014b), the high numbers in our study is attributed to less rigor in its prediction. Among di- and tri-nucleotide repeats identified, AG/CT, CGC & AAG repeats were most prevalent, which was in accordance with earlier reports (Sood et al., 2016).

One hundred informative primers were synthesized and standardized in *E. coracana* species. Although 88 primers showed amplification, only 60 primers showed clear and scorable alleles. From this 60 primers, a set of 38 primers was selected for genotyping 23 landraces and 9 cultivated varieties (Table 1). Twenty eight primers were polymorphic that amplified a total of 91 alleles with the mean values of 4.6, 0.5 and 0.43 corresponding to the number of alleles per locus, diversity index and PIC (Table 2). Earlier studies in finger millet indicate the average polymorphisms of 72 % has been reported while characterizing 64 Indians and 61 non-Indian finger millet genotypes using 87 genomic SSRs (Ramakrishnan et al., 2016). While EST-derived SSRs exhibit low polymorphism, our study shows up to 55 % polymorphism, which is higher than the recently reported 48 % in finger millet (Babu et al., 2014b). Majority of the primers were multi-allelic that amplified a maximum of up to 12 alleles (Primer ECML036), with an average of 4.6 alleles per locus. Seven primers amplified two loci each per locus which is predictable as the cultivated *E. coracana* is a tetraploid species derived from two genomes (Goron and Raizada, 2015). A goodness-of-fit test for HWE indicated 7 loci deviating significantly from expectations at $P \leq 0.001$ (Bonferroni correction) as the population size was small which could be improved by additional tests with large populations of unrelated individuals. The gene ontology (GO) terms retrieved from Uniprot database classified the functions into 6 major groups in the categories of molecular function, biological process and cellular component. Majority the sequences were homologous to the genes responsible for the enzyme catalytic activity that are involved in biotic and abiotic stress tolerance (Figure 1). Similarly, the functional annotation of finger millet sequences classified 50 % of the genes to catalytic activity in GO terms (Gimode et al., 2016). Functional annotation and classification of markers becomes essential as functional SSRs could be helpful in comparative mapping and post transcriptional gene silencing (Nagabhushana et al., 2017).

Table 2: Details of Twenty Microsatellite Primers Used for Characterization of 23 Landraces and 9 Varieties of Finger Millet

Primer ID	GenBank ID	Repeat	Forward Primer(5'-3')	Reverse Primer (5'-3')	T _A (°C)	Size (bp)	k	He	PIC
ECML001	EU439937.1	(CGC) ₃ *	CTGGATCATGCACGAGTACC	CTTCCTCCTCCTTTCAGTG	58	180	6	0.57	0.46
ECML004	BV723953.1	(TG) ₆ ttg	TGCATGCAAGCTTTCCTAT	ACACCACCAACCCACACA	58	260	8	0.83	0.79
ECML005	BV723953.1	(GCGT) ₄	CTAGTAGCAGAAATCACGCC	TGTCATCATCACTCGCATGG	58	223	3	0.40	0.33
ECML006	BV723951.1	(CAA) ₅ cac	CCTCCCTAGCAGAATCAGGT	TCAGCCTTATACACCGTTGC	58	164	5	0.57	0.49
ECML007	BV723950.1	(AC) ₁₁ (AG) ₂₃	TTGCTTCTGATGCATCTCCC	CCCTGACACAGAAACAGGAC	58	257	6	0.52	0.45
ECML010	BV723947.1	(CT) ₁₀ gt	GGATCCAAACACCCCATCAA	TTGTCGTGTAAGCTTTGCT	58	173	9	0.64	0.57
ECML011	BV723946.1	(TTG) ₅ tg(GTT) ₇ gg	CTTTGGTGCTGGTGATAGGG	TTCATCTCAGCAGCAGCAAA	58	274	5	0.72	0.67
ECML015	BV723943.1	(TTG) ₄ ctg	GAGGATGACTCCATTGCTG	CTCTCGCCACCATTAGTAGA	58	209	4	0.59	0.51
ECML021	BV723932.1	(TA) ₅ att	GCGTAAGTCTTGCAGGATAT	CCCTCACCTCTATGGTAG	58	277	2	0.45	0.35
ECML024	GF107744.1	(GAT) ₇	AGGTGACCTCAAGACCAAGA	TGGATCCACAGTTGAAGCAC	58	282	4	0.60	0.53
ECML029	GF107728.1	(AG) ₈ gat	GCCGCGAGCACTAATTTTITA	GGCAAAAGATACTCTCTCTCCC	57	286	2	0.22	0.20
ECML036	GF107710.1	(GT) ₂₉	AAAGGTTGTGAGGGAGGAGA	CAGAACAAACAGGACCCACAA	58	297	12	0.84	0.81
ECML051	JN378851.1	(GCC) ₃ atc	ATTGCGTTCGACGATGCAA	GTGTTTGTGAGCTGGAGAA	58	275	2	0.22	0.20
ECML052	KC894819.1	(GT) ₃₇	GCATATGTGTTGATGTCGGC	CCCCAAGTTCGGATCAAC	58	178	2	0.22	0.20
ECML071	GF107753.1	(GA) ₂₁	CTCGTAGCAGAATCACAGG	CTGCCACGAGGATAGATGC	58	255	3	0.67	0.58
ECML078	EB187460.1	(AGC) ₇	GCCAAAGGAAACTTTGAGGC	AGGCCGAGAGAAAGAGAGAG	58	211	4	0.50	0.37
ECML079	EB187410.1	(TTC) ₈	CGGCAATGCCATCATTTTGT	ACGCCAACAATATCGAACCA	58	165	4	0.51	0.37
ECML080	EB086245.1	(TC) ₁₅	GGGACTCTAGTTTCCGCTTT	GTTTGGGGAAGAAGCTCCAA	57	235	2	0.26	0.22
ECML086	AF144580.1	(GAA) ₃ atg	CTGGAACCTTTCTCCAACGA	ACCCGAGAAAATCGAAGCAA	57	285	2	0.22	0.20
ECML091	AY523508.1	(TCAAT) ₂ tct	GGGGGAGAGAAAACTTTTGG	ACAATCAAGTCCGTAGCGTC	57	264	6	0.46	0.39
Mean							4.6	0.50	0.43

Note: T_A Annealing temperature; k-Number of alleles amplified; He-Expected heterozygosity; PIC-Polymorphic Information Content. The He values in bold numbers indicate markers deviating from Hardy-Weinberg equilibrium (p=0.001)

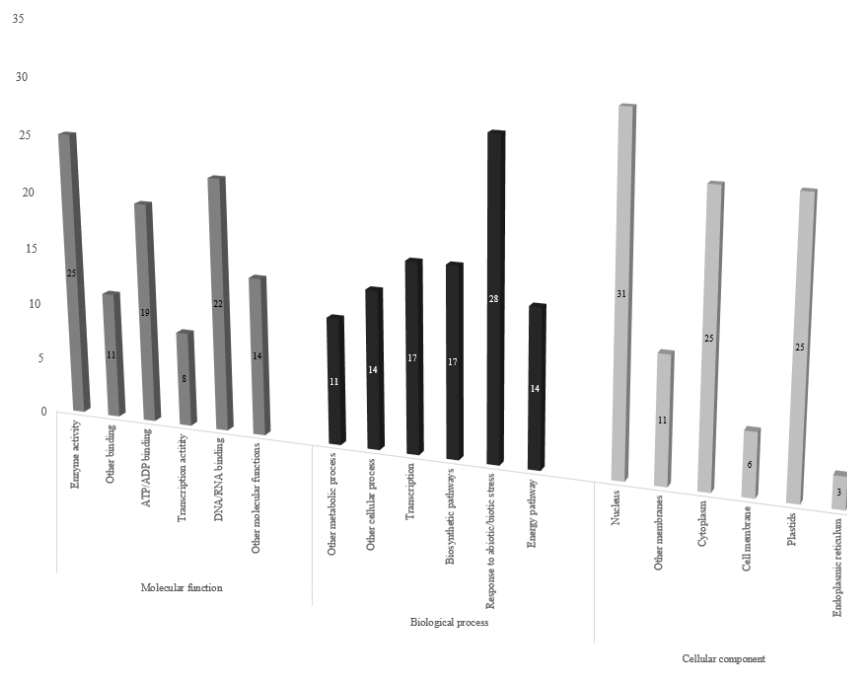


Figure 1: Gene Ontology Based Classification of Markers under Different Functional Categories Viz., Molecular Function, Biological Process and Cellular Component. The Numbers Embedded in the Graphs Indicate Percent Distribution under Each Group

The AMOVA results indicated the prevalence of significant variation within (80%) and among (20%) populations, indicating the scope for exploiting the genetic diversity for further selection and hybridization programs. Similar results of high variation within populations were reported by Ramakrishnan et al., (2016) and Babu et al., (2014b). The PCoA analysis drawn from 1st and 2nd coordinates based on highest Eigen values indicated wide variation spread

across the genotypes from Shimoga (Malnad region), whereas the individuals from other locations were more or less clustered together in 4 different groups (Figure 2). As the millets are grown in remote regions, significant diversity persists in the local landmarks that are valued for a wide variety of traits.

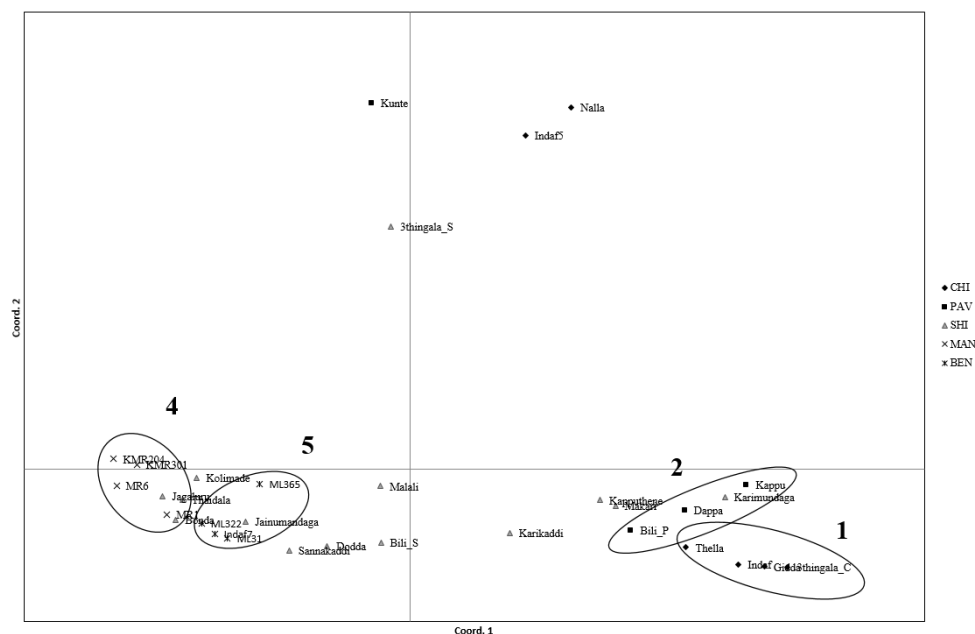


Figure 2: Principal Co-Ordinate Scatter Diagram of 23 Landraces and 9 Varieties of Finger Millet Collected from 5 Locations (CHI: Chintamani, PAV: Pavagada, SHI: Shimoga, MAN: Mandya BEN, Bengaluru). The Numbers in the Diagram Indicate Different Clusters Based on Respective Locations

The population structure analysis was performed with five independent runs of $K = 1$ to 10 at 5000 Markov chain Monte Carlo (MCMC) repetitions and 50000 burn-in period, using prior population information with the admixture model assuming correlated allele frequencies. Furthermore, ΔK values were estimated based on the rate of change in the log probability of data between successive K values and Evanno's correction (Evanno et al., 2005) showed two peaks of ΔK , for $K = 1$ & 4, suggesting the presence of four subgroups within a major group (Figure 3 & 4). Except few admixtures, the grouping of individuals under four subgroups followed their geographical location and these results were in correspondence with PCoA results. The individuals from Bengaluru and Mandya were genetically closer than the rest of the individuals, as they were all developed varieties. With the limited scope for hybridization between wild species and cultivated varieties, the landraces with ample of diversity offer an excellent source of genetic diversity in crop development (Goron and Raizada, 2015). Although genomic microsatellite markers are known for their abundance in the genome and high polymorphism, EST-derived microsatellites are considered more informative due to their association with functional genes and transferability across related species and genera. The association of functionally derived SSR polymorphisms with actual traits have been demonstrated in finger millet aiding in cost-effective marker assisted selection programs for crop improvement.

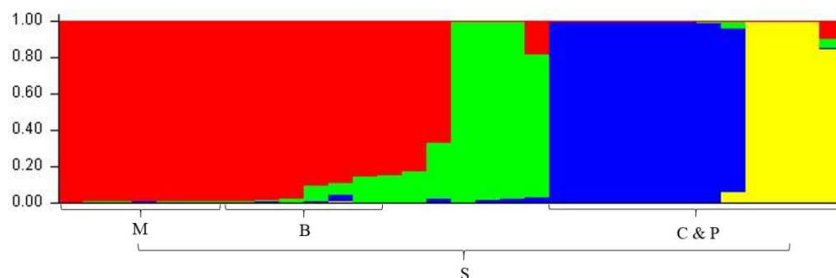


Figure 3: Population Structure of 23 Landraces and 9 Varieties of Finger Millet. The Different Colors Indicate Four Subgroups with Admixture Alleles among Genotypes from 5 Locations Viz., C: Chintamani, P: Pavagada, S: Shimoga, M: Mandya, B: Bengaluru

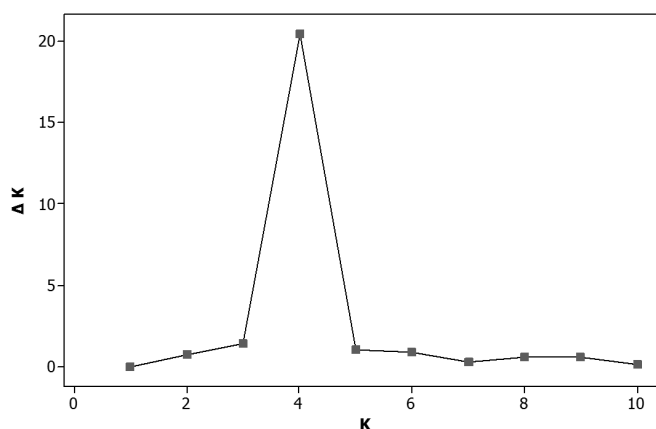


Figure 4: Structure Estimation of the Number of Populations for K Ranging from 1 To 10 by ΔK Values. the Value of K was Significant At K = 4

CONCLUSIONS

Even with the recent development of increased genomic resources in finger millet, wide phenotypic diversity is not reflected at genetic level due to low genetic polymorphism, which has posed a major challenge in molecular mapping approaches. Unlike in major crops, the availability of informative genomic tools and its subsequent utility in characterization of wild cultivars and landraces is still a major concern in finger millet, which requires the attention of researchers from India and Africa. From this perspective, the current study highlights large scale development of EST-derived SSRs and their application in characterizing local landmarks. Therefore, the narrow genetic diversity existing within cultivated finger millet can be addressed by utilizing the local landmarks that hold considerable genetic diversity. These markers can be successfully utilized in genetic characterization; population studies, genome mapping and marker assisted breeding programs.

ACKNOWLEDGEMENTS

The authors thank Dr. Zaheer Basha, University of Agricultural Sciences, Bengaluru and Dr. Gnanesh, University of Agricultural Sciences, Shimoga for the constant support in arranging the plant material. Sincere thanks extended to Mr., Suresh Kumar for the Bioinformatics support in SSR discovery.

REFERENCES

1. Babu, B.K., Agrawal, P.K., Pandey, D., Jaiswal, J. P. & Kumar, A. (2014a). Association mapping of agro-morphological characters among the global collection of finger millet genotypes using genomic SSR markers. *Mol. Biol. Rep.* 41, 5287-5297
2. Babu, B. K., Agrawal, P. K., Pandey, D. & Kumar, A. (2014b). Comparative genomics and association mapping approaches for opaque2 modifier genes in finger millet accessions using genic, genomic and candidate gene-based simple sequence repeat markers. *Mol. Breed.*, 34, 1261-1279
3. Babu, B.K., Dinesh, P., Agrawal, P.K., Sood, S., Chandrashekara, C., Bhatt, J. C., et al. (2014c). Comparative genomics and association mapping approaches for blast resistant genes in finger millet using SSRs. *PLoS ONE*, 9, e99182
4. Babu, B.K., Pandey, D., Agrawal, P.K., Sood, S., & Kumar, A. (2014d). In-silico mining type and frequency analysis of genic microsatellites of finger millet (*Eleusine coracana* (L.) Gaertn.): a comparative genomic analysis of NBS-LRR regions of finger millet with rice. *Mol. Biol. Rep.* 41, 3081-3090
5. Chandra, D., Chandra, S., Pallavi & Sharma, A. K. (2016). Review of Finger millet (*Eleusine coracana* (L.) Gaertn.): A powerhouse of health benefiting nutrients. *Food Science and Human Wellness*, 5, 149-155
6. Doyle, J. J. & Doyle, J. L. (1990) Isolation of plant DNA from fresh tissue. *Focus*, 12, 13-15
7. Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.*, 14, 2611-2620
8. Gimode, D., Odeny, D. A., deVilliers, E. P., Wanyonyi, S., Dida, M. M., Mneney, E. E., Muchugi, A., Machuka, J. & deVilliers, S. M. (2016). Identification of SNP and SSR markers in finger millet using next generation sequencing technologies. *PLoS ONE*, 11(7), e0159437
9. Goron, T. L. & Raizada, M. N. (2015). Genetic diversity and genomic resources available for the small millet crops to accelerate a New Green Revolution. *Front. Plant Sci.*, 6, 157
10. Kalinowski, S.T., Taper, M.L. & Marshall, T.C. (2007). Revising how the computer program CERVUS accommodates genotyping error increases success in paternity assignment. *Mol. Ecol.*, 16, 1099-1106
11. Nagabhushana, K., Prasad, S.H. & Rajkumar, R. (2017). Intra and intergeneric transferable gene-derived orthologous microsatellite markers in *Eucalyptus* and *Corymbia* species. *Journal of Forest Research*, 22 (1), 65-68
12. Peakall, R. & Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research-an update. *Bioinformatics*, 28, 2537-2539
13. Pritchard, J. K., Stephens, M. & Donnelly, P. (2000a). Inference of population structure using multilocus genotype data. *Genetics*, 155, 945-959
14. Ramakrishnan, M., Antony, C. S., Duraipandiyan, V., Al-Dhabi, N.A. & Ignacimuthu. (2016). Assessment of genetic diversity, population structure and relationships in Indian and non-Indian genotypes of finger millet (*Eleusine coracana* (L.) Gaertn.) using genomic SSR markers. *SpringerPlus*, 5, 120
15. Salimath, S. S., Oliveira, A. C. D., Godwin, I. D. & Bennetzen, J. L. (1995). Assessment of genome origins and genetic diversity in the genus *Eleusine* with DNA markers. *Genome* 38, 757-763
16. Sood, S., Kumar, A., Babu B. K., Gaur, V. S., Pandey, D., Kant, L. & Pattnayak, A. (2016). Gene Discovery and Advances in Finger Millet [*Eleusine coracana* (L.) Gaertn.] *Genomics-An Important Nutri-Cereal of Future*. *Front. Plant Sci*, 7, 1634
17. Thiel, T., Michalek, W., Varshney, R. K. & Graner, A. (2003). Exploiting EST databases for the development and

- characterization of gene-derived SSR-markers in barley (Hordeumvulgare L.) TheorAppl Genet.* 106 (3), 411-22
18. Untergrasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B. C., Remm, M.&Rozen, S. G. (2012).Primer3 - new capabilities and interfaces. *Nucleic Acids Research*,40 (15), e115
19. Yadav, S., Gaur, V. S., Jaiswal, J. P.&Kumar, A. (2014b). Simple sequence repeat (SSR) analysis in relation to calcium transport and signalling genes reveals transferability among grasses and a conserved behaviour within finger millet genotypes. *Plant Syst. Evol.*, 300, 1561-1568